

Process-oriented social responsibility indicator for mutual funds: A multi-criteria decision analysis approach

Tim Verheyden¹

PhD researcher

KU Leuven

Lieven De Moor²

Associate professor

Vrije Universiteit Brussel

Keywords: socially responsible investing; mutual funds; multi-criteria decision analysis.

JEL-codes: A13, C44, G10, G11, G12

Word count article: 7,993 words

¹ Corresponding author, Katholieke Universiteit Leuven (KULeuven), Campus Brussels (HUBrussel), Faculty of Economics and Business, Warmoesberg 26, 1000 Brussels, Belgium, tim.verheyden@kuleuven.be, +32472-38 39 09 (tel.)

² Vrije Universiteit Brussel (VUB), Faculty of Economic and Social Sciences and Solvay Business School, Pleinlaan 2, 1050 Brussels, Belgium, lieven.de.moor@vub.ac.be, +32498-616463 (tel.)

Process-oriented social responsibility indicator for mutual funds: A multi-criteria decision analysis approach

Abstract

In recent years, socially responsible investing (SRI) has grown from a niche to a mainstream investment strategy. Consequently, the supply of SRI mutual funds expanded and a discussion on the performance of SRI funds developed. The traditional approach to measuring mutual fund performance does not take into account possible heterogeneity with respect to social responsibility and focuses on a dichotomous SRI versus non-SRI typology. We consider multi-criteria decision analysis (MCDA) to build an indicator to capture the multi-faceted nature of SRI and to help control for differences in process-oriented social responsibility. From the wide variety of established MCDA methods, we determine and implement four eligible techniques. Considering extensive robustness analysis and additional criteria, we find the PROMETHEE methodology to be most recommended for both academic and professional applications. The obtained scores and rankings prove to be highly robust and are in line with generally accepted intuition in the industry. The proposed PROMETHEE-based indicator is also easy to understand and implement, which makes it highly suited for implementation in mutual fund performance research (academic application). The available range of sensitivity analyses in the PROMETHEE technique is also instrumental to help fund managers optimize the design of new mutual funds with respect to social responsibility (professional application).

1. Introduction

The practice of socially responsible investing (SRI) has evolved over the past couple of decades. Starting from the moral investing principles of religious organizations, the first generation of SRI was only concerned with excluding companies that did not meet certain ethical standards from the investment universe. This practice is referred to as negative screening. In a later generation, SRI shifted its focus to the application of positive screens, i.e. companies that were performing

well according to environmental, social and governance (ESG) issues were selected from the investment universe. The third generation of SRI simply combined the use of positive and negative screens. Shareholder activism is the latest addition to this set of SRI practices. Eurosif (2010) segments the market by means of core versus broad SRI. Core SRI consists of exclusions based on norms and ethical values, and positive screening (incl. best-in-class approach). Broad SRI is composed of simple screening strategies, engagement and integration. The latter category represents the mainstreaming of SRI.

Following several societal challenges like the Vietnam War, the civil rights movement and the racist system of apartheid, the interest in SRI has grown over the years. In the United States, about one out of every nine dollars under professional asset management is now invested in SRI assets (Forum for Sustainable and Responsible Investment, 2012). The volume growth of SRI is also larger than for the conventional asset class. In Europe, we see similar results, with the combined growth of SRI strategies outperforming the conventional market (Eurosif, 2012).

Despite the growth in the SRI market, and the associated maturation from a niche to a mainstream strategy (Sparkes & Cowton, 2004; KPMG & ALFI, 2013), there is no conclusive evidence of SRI funds not underperforming conventional funds. In the field of strategic management, the debate goes back to the opposing views of Friedman (1970) and Freeman (1984), i.e. a shareholder versus a stakeholder view of the firm. Mackey, Mackey, and Barney (2007) found that the effectiveness of social responsibility crucially depends on the demand and supply conditions for SRI: a finding that reconciles between both extreme views. In financial economics, the literature goes back to the opposing views of Moskowitz (1972) and Markowitz (1952). Moskowitz (1972) published the first empirical paper that indicates there can be financial value to screening strategies. Markowitz (1952), the creator of modern portfolio theory, argues that SRI funds are a subset of the market portfolio and suffer from idiosyncratic risk because of the fewer diversification opportunities that come with the additional constraints to the investment universe. Today, scholars mostly apply asset-pricing regression models to compare risk-adjusted returns between SRI and conventional mutual funds. Notwithstanding the increasing complexity of these methods, results found are often inconclusive and/or weak. The typical approach is to use a multi-factor asset-pricing regression accounting for different sources of risk. Next, the

model is applied to a difference portfolio of both SRI and conventional funds, which yields a risk-adjusted return (alpha; constant in the regression). If the constant is significantly different from zero, a conclusion can be drawn on the out- or under performance of SRI funds.

A main drawback of earlier research is that only typical sources of risk are taken into account, without considering the possible impact of style differences between and among SRI and conventional funds, with respect to the process quality of designing funds based on ESG criteria. Renneboog, Ter Horst and Zhang (2008a) attempt to overcome this issue using an “ethics factor”, but do not take into account the multi-dimensional nature of social responsibility. Therefore, the risk-adjusted difference returns found in previous research might be biased. It is reasonable to assume that social responsibility can affect the long-term risk faced by corporations. For example, when a firm is investing in renewable energy, it is less likely to be negatively impacted by an energy crisis in the future. A better relationship with the employees can help reduce the chance of future strikes. Social responsibility should thus be taken into account as a factor when examining the performance of mutual funds. Considering the specific social responsibility content of individual firms is unfeasible at the aggregate level of a mutual fund. Alternatively, we build a proxy indicator for social responsibility by measuring the quality of the design process of putting together SRI mutual funds. Such a SRI process quality indicator can be used to help discriminate between mutual funds with respect to process-oriented social responsibility, instead of the traditional dichotomous approach (SRI vs. non-SRI), which does not allow researchers to account for heterogeneity among both the group of SRI and conventional funds. This can add more detailed insights into the risk-adjusted performance relation between the different types of funds. Next to the academic application, these scores can be insightful to address professional challenges as well. For example, managers could use such a tool when constructing new investment funds.

Different studies have tried to measure a mutual fund’s social responsibility in the past. Most commonly, several papers have described the development of tools for individual investors to take into account both financial and ESG perspectives. Hallerbach, Ning, Soppe and Spronk (2004) present an interactive multiple goal programming approach to help investors compose SRI portfolios from individual assets. Cabello, Ruiz, Pérez-Gladish and Méndez-Rodríguez

(2014) complement this research by proposing a reference point method for scoring SRI mutual funds with respect to environmental performance. Despite its usefulness for individual investors in making better investment decisions, these kinds of approaches are rendered ineffective for implementation in further SRI performance research given the required number of subjective assessments by the decision maker. These approaches are also less instrumental for fund managers looking to optimize the design of new funds. Rather than individual investor decisions, we focus on SRI performance research and possibilities to take into account heterogeneity among funds with respect to ESG criteria. Additionally, we consider to what extent this kind of method can be used in a more professional setting to improve the design of mutual funds with respect to social responsibility. Ideally, this calls for techniques that are easy to implement and replicate and thus do not require any input from experts or individuals to assess criteria or elicit preferences. One way to achieve this goal is to simply consider the number and variety of applied screens to approximate social responsibility (e.g. Renneboog, Ter Horst & Zhang, 2008b). However, this approach reduces the concept of SRI to merely screening activities. To build a process quality indicator that reflects the multiple dimensions to the concept of SRI, we therefore consider the field of multi-criteria decision analysis (MCDA). Pérez-Gladish and M'Zali (2010) and De Moor, Devooght and De Bondt (2012) were the first to design an indicator to measure the process-oriented social responsibility of mutual funds. However, they only implemented the AHP methodology and disregarded other possibly superior methodological approaches. In contrast, we consider the entire set of established MCDA techniques. From an extensive analysis of the different methods within the MCDA framework, four eligible methods are applied. To find which of the four considered MCDA methods performs best to approximate social responsibility in a way that can be useful for future SRI performance research, we conduct extensive robustness analysis and take into account additional criteria depending on the type of application.

2. MCDA methodologies

2.1. Background theory of MCDA

Multi-criteria decision analysis (MCDA) is an operations research tool to help solve delicate decision problems that involve multiple conflicting decision factors (Figueira, Greco, & Ehrgott,

2005). Within the MCDA paradigm, there often is no such thing as the optimal solution. Rather, a decision maker evaluates a set of alternatives considering all decision factors or criteria that are relevant to the analysis. The main difficulty in the decision process is the aggregation of relevant criteria. The specificity of the different methods within MCDA lies mostly in the way this aggregation problem is addressed. Given the multiple dimensions involved when assessing the social responsibility of mutual funds, and the need for subjective judgments, MCDA provides an interesting framework to build an SRI design process quality indicator. The transparency of the MCDA framework can help decision makers (e.g. fund managers) to better understand how the social performance scores were calculated. All in all, the MCDA tools can help to objectivize some of the subjective information when assessing social responsibility.

Every MCDA method consists of five building blocks (Belton & Stewart, 2002). In a first step, a decision maker determines the consistent family of relevant criteria (Roy, 1974). Next, alternatives considered to the decision problem are introduced. Then, every criterion is scored with respect to each alternative to construct a performance table. To determine the relative importance of the different criteria, weights need to be elicited. This can be done in a number of ways (e.g. directly or via pairwise comparison). Finally, the performance table is transformed into a score using approaches that are specific to the different MCDA methods.

The simplest approach to address multi-criteria problems is a weighted sum (Figueira et al., 2005). However, this requires the often-untenable assumption of linearity of preferences. To better accommodate complex decision processes, a plethora of methods has been developed in the MCDA field. These methods can be divided into three schools of thought: value measurement methods, outranking methods and goal, aspiration or reference level methods. Assuming that a bad score on one criterion can be compensated by a good score on another criterion, value measurement methods evaluate a global score for every alternative based on partial scores with respect to the different criteria. Outranking methods allow for the notion of incomparability, i.e. partial scores cannot be compensated, and are based on pairwise comparisons. Using feasibility constraints, goal, aspiration or reference level methods can allocate goals to every criterion to identify the most desirable alternative.

For the purpose of our process quality indicator, we examine the established MCDA methods. To find which methods are feasible to our decision problem, we consider different criteria that apply to our ideal process quality indicator. First, we require a quantitative output for the indicator, as we want to use the score in asset pricing regressions. The method also needs to be state-of-the-art. Additionally, we need to be able to deal with qualitative inputs, not just hard quantitative inputs. The criteria used in assessing the process quality should be easy to translate into the model. Furthermore, the method is ideally easy to implement, using either existing software or by programming in statistical software. Finally, we want to limit the amount of expert judgments required to calculate the scores. The comparison of the different MCDA methods with respect to these criteria can be found in Table 1.

Comparisons	Score output	State-of-the-art	Qualitative input	Translation of criteria	Implementation	Judgments
<i>Value measurement methods</i>						
Analytic hierarchy process	x	x	x	x	x	x
Analytic network process	x	x	x	x	x	
MACBETH	x	x	x	x	x	x
Multi-attribute utility theory	x	x	x	x	x	x
UTADIS/GRIP	x	x	x	x	x	
<i>Outranking methods</i>						
ELECTRE			x	x		
PROMETHEE	x	x	x	x	x	x
<i>Goal, aspiration or reference level methods</i>						
TOPSIS	x	x	x	x	x	x
Goal programming	x	x			x	x
Data envelopment analysis*	x	x			x	x
<i>Other</i>						
Decision rules		x	x	x	x	

Table 1: MCDA method comparison table

The different methods from the field of MCDA assessed with respect to the ideal characteristics of the process-oriented social responsibility indicator.

Note that we also included data envelopment analysis (DEA) as a possible method, even though it does not really belong in the MCDA universe. From the analysis, we find five methods to be highly feasible: analytic hierarchy process (AHP), MACBETH, multi-attribute utility theory (MAUT), PROMETHEE and TOPSIS. Multi-attribute utility theory and PROMETHEE normally call for elaborate value judgments, but the specific definition of our decision criteria makes them feasible to our problem as well (see section 2.2). The value measurement methods seem most appropriate to build a process quality indicator. Specifically, we consider the analytic hierarchy process (AHP), MACBETH and multi-attribute utility theory (MAUT). From the outranking school, we implement the PROMETHEE II ranking methodology, which yields a

complete ranking of alternatives based on net flow scores, which can be used as a quantitative input in asset pricing regressions to control for differences in social responsibility. Goal, aspiration or reference level models are generally less well suited as most of the methods call for highly quantifiable criteria and are mostly useful when the set of alternatives is continuous and infinite rather than discrete and limited. One exception is the simple TOPSIS methodology, which we apply as well. From a thorough robustness analysis and additional criteria depending on the type of application of the scores, we are able to determine which specific MCDA method is most recommended to serve as a design process quality indicator in future research.

Before introducing the different MCDA methods we implement to find a social responsibility proxy score, we first present two building blocks needed in every MCDA exercise: the relevant criteria, and the decision alternatives considered.

2.2. Criteria

In building an MCDA indicator, we need to define assessment criteria. Following Roy (1974), a consistent family of criteria satisfies the axioms of exhaustibility, cohesion and non-redundancy. Pérez-Gladish and M'Zali (2010) conduct extensive literature research to come up with a hierarchy that is split up between criteria associated to the actual content of a SRI fund on the one hand, and the transparency and credibility of the SRI fund on the other hand. A major drawback of their study is the need for expert judgment in comparing different alternatives, which makes it harder to replicate the methodology in other research. De Moor et al. (2012) use the same split of criteria, but formulate them in such a way that information can simply be extracted from publicly available information about the fund (e.g. prospectus), without need for extra judgment from the decision maker, who now only needs to provide preference information about the importance of criteria. As we aim to build a process quality indicator that can easily be implemented in future financial economics research, we start from the criteria hierarchy of De Moor et al. (2012). Then we omit three criteria³ because of redundancy, generalize another

³ Assessment by means of controversies, exclusion of disputable technologies and active and regular audit of investments.

criterion⁴ and add one complementary criterion⁵ that was pointed out by Pérez-Gladish and M'Zali (2010), the Belgian financial sector organization Febelfin (2012) and the United Nations (2013) responsible investment initiative.

The eventual hierarchy (Table 2) consists of different statements about the process-oriented social responsibility of a mutual fund. Note that all of these criteria concern the social responsibility content of the design process of how a mutual fund was constructed, rather than the company-specific social responsibility content. Hence, we refer to our indicator as process-oriented. Either a fund complies with the criterion or it does not. As a consequence, the performance table is dummy-coded and only requires publicly disclosed information without need for any expert judgment. Even though this constraint on the performance table limits the power of several MCDA methods to some extent, this approach towards the criteria greatly contributes to the replicability of the indicator for future research. An opposite approach, which allows for more elaborate judgment, can lead to richer results, but would not be feasible for implementation in asset pricing research. As our eventual goal is to shed new light on the performance relation between SRI and conventional funds, we opt for the restrictive approach. The dummy coding also simplifies the process of eliciting value functions, which solves the need for more elaborated judgment in multi-attribute utility theory and PROMETHEE. This in turn means that we have five feasible methods to build a process indicator. Because of the aggregation and comparison across 20 different criteria, the scoring output is not dichotomous, which still enables us to capture the heterogeneity between different mutual funds with respect to the quality of the SRI screening process.

⁴ Investment is principally (> 75%) in companies that invest in sustainable technologies (not just sustainable electricity and CO₂ reduction).

⁵ Incorporation of SRI principles established by reputable organizations.

Goal	Criteria	Subcriteria	Sub-subcriteria
Social performance indicator	Screening process and consistency	Priority screening process	
		Independent data gathering and analysis of sustainability	Data gathering and analysis of sustainability by independent external specialists (e.g. EIRIS)
			Incorporation of SRI principles established by reputable organizations (e.g. UN SRI, Febelfin)
			Information from stakeholders and relevant NGOs
		Positive selection criteria	Best-in-class approach for criteria with respect to ESG criteria
			Use of sector specific positive criteria
			Investment is principally (> 75%) in companies that invest in sustainable technologies
		Negative selection criteria	Use of categorical rejects
			Assessment by means of negative criteria
			Exclusion of unsustainable technologies
		Monitoring and updates	A research teams checks legal and regulatory developments, trends and behavior of companies such that criteria are in line with recent societal developments
			Monitoring if portfolio is consistent with defined criteria (continuously, sector specific or occasion specific)
		Dialogue	Companies are informed about conclusions selection methodology
			Active engagement policy (constructive and critical dialogue with companies in portfolio)
			Active voting policy (voting at companies' shareholder meetings)
	Transparency and control	Transparency	Release of qualitative information about the screening process (e.g. applied screens)
			Release of quantitative information about the screening process (e.g. scores)
			Release of current portfolio
			Compliance with external transparency guidelines (e.g. Eurosif/Belsif)
		Board of experts	

Table 2: Hierarchy of criteria

Schematic presentation of the hierarchy of criteria, with the 20 bottom-level criteria used in the analysis. An in-depth presentation of the different criteria can be found in Appendix A.

2.3. Alternatives

Next to the criteria, every MCDA method also calls for specific alternatives to be defined. In the case of a SRI process indicator, the ideal set of alternatives consists of every mutual fund available to an investor. However, it would be unfeasible to gather the necessary data to apply the MCDA methods to such a large set of alternatives. Instead, we opt for a small sample set of alternatives, as the focus of this paper is on the indicator methodology. We simply illustrate the application of the indicator to a limited set of alternatives from SRI mutual fund providers on the Belgian market. The choice for the Belgian market is motivated by the availability of an official list of recognized sustainable products⁶ by the Belgian financial industry organization (Febelfin), which is accompanied by additional documentation on the mutual fund design process for the different SRI providers active on the Belgian market. This kind of information simplifies the completion of the performance table to a great extent. Specifically, we consider 9 alternatives, which can be divided into two categories according to whether the fund is labeled as an SRI fund or as a conventional fund by its issuer (Table 3).

⁶ <http://sustainableproducts.febelfin.be>

SRI funds	Conventional funds
BNPP L1 Equity World Aqua (LU0831546592)	BNPP L1 Equity World (LU0072778490)
Dexia L Sustainable World (LU0113400328)	Dexia Quant Equities World (LU0235267860)
ING L Invest Sustainable Equity (LU0394658412)	ING L Invest World (LU0119219730)
KBC EcoFund World (BE0133741752)	KBC Equity Fund World (BE6213775529)
Triodos Sustainable Pioneer Fund (LU0278272843)	

Table 3: Alternatives

Presentation of the funds considered in the analysis, including the respective ISIN-code.

For the four largest commercial banks active in Belgium, we include both a SRI fund and a matching (e.g. same international orientation) conventional fund. Additionally, we include a fund from Triodos, which is a commercial bank that specifically focuses on sustainable finance. All of the funds are capitalization funds, are geared towards equity and are oriented towards the global market.

Future research should look into the further application of the scoring tool on more and other alternatives. The only element of the analysis that changes is the performance table. To limit the data gathering effort, data envelopment analysis (DEA) could also be applied a priori to a broader set of alternatives to bring it down to a more limited set that is more feasible for the MCDA indicator.

2.4. Feasible methodologies

After establishing criteria and alternatives, we now need to find an appropriate aggregation method to calculate an overall SRI process quality score for each mutual fund. We specifically aim for a scoring output, as we want to include this score in future SRI performance research. For other purposes, like the development of an SRI tool for regulatory organizations, it might be wiser to search for a classification, rather than a scoring/ranking output.

The different MCDA methods use alternative approaches to address the aggregation problem. We consider five methods, which are suited for the SRI performance indicator problem: analytic

hierarchy process (AHP), MACBETH, multi-attribute utility theory (MAUT), PROMETHEE and TOPSIS. Next, we describe each of these five methods, why they are useful for the purpose of a process quality indicator and how they are implemented.

2.4.1. Analytic hierarchy process

First presented by Saaty (1977, 1980), the analytic hierarchy process (AHP) is an elegant approach for dealing with discrete alternative problems with multiple conflicting criteria. The technique can be classified as a value measurement method. In essence, AHP consists of three steps. In a first step, the problem needs to be structured by building a hierarchy of criteria relevant to the nature of the problem at hand. On top of the hierarchy, the goal of the AHP analysis is stated. Below the goal, criteria are listed. Depending on the problem, it might be necessary to create additional levels of sub-criteria and sub-sub-criteria as well. The final level in the hierarchy consists of the alternatives considered. An important assumption for each of the methods we implement is the independence of criteria. The AHP hierarchy for the SRI process indicator looks like the criteria in Table 2, with the addition of the alternatives at the end of each of the 20 bottom-level criteria. The same hierarchy is also used for the other methods implemented in this paper.

The second step of the AHP is the scoring process, which leads to a priority ranking of alternatives and criteria. Alternatives need to be scored with respect to all the lowest-level criteria in the hierarchy. In our case, we define the criteria in such a way that alternatives receive a true (1) or false (0) quote for every criterion, as was explained before. For example, if we find in a fund's publicly disclosed information that the screening process is given priority over the financial analysis, the fund receives a true (1) score for that criterion. The eventual performance table can be found in Appendix B. Next to scoring the alternatives with respect to the criteria, priorities need to be calculated. In the AHP this is typically done by pairwise comparison of the bottom-level criteria within the same group using a verbal importance scale from 1 to 9. Psychologists argue that this technique is easier for decision makers when expressing their preferences with regard to the importance of different criteria. Instead of allocating global weights to the criteria in the hierarchy, AHP guides the decision maker through several pairwise

comparisons of bottom-level criteria from the same subgroup, indicating which criterion is preferred and to what extent. All of these pairwise comparisons are then combined into a comparison table, which is then checked for consistency. When the table is deemed consistent the criteria priorities and overall priority scores can be used in SRI performance research. An overview of the questionnaire including the AHP comparisons presented to the two experts to elicit weights can be found in the online appendix.

In a last step of the AHP, a robustness analysis is usually conducted to examine how much the results change when inputs (e.g. weights) are changed. This is an important step to determine the reliability of results. Note that the results from one of the experts are used to conduct our initial analysis. Results from the second expert are used to test for overall robustness, in addition to the analysis of stability intervals. To implement the AHP technique in building an SRI process indicator, we make use of the Expert Choice software package.

2.4.2. MACBETH

MACBETH, which stands for “Measuring Attractiveness by a Categorical Based Evaluation Technique”, is very similar to the AHP and can also be considered a value measurement approach (Bana e Costa, De Corte, & Vansnick, 2005). The first step again requires a hierarchy of criteria. We use the same hierarchy as for the AHP (Table 2).

In the second step, the scoring process needs to be completed. Like for the AHP, the alternatives can be evaluated directly for each criterion, using publicly available information about a fund. MACBETH also requires the decision maker to indicate the appreciation for an increase from false (0) to true (1) for every criterion using a seven-point semantic scale. Next, criteria are weighted using pairwise comparisons. The difference with AHP, however, is in the scale used to make the comparisons. Instead of a ratio scale, MACBETH uses an interval scale. In making the comparisons, MACBETH requires the decision maker to indicate preferences on a semantic scale with seven categories. Unlike AHP, MACBETH also requires the pairwise comparison of every bottom-level criterion, and not just those bottom-level criteria within the same subcategory. For the process quality indicator, this greatly increases the number of judgments

needed, which adds complexity to the assessment. Instead, we simply consider the weights found by means of the AHP approach. This should not depreciate the value of the overall findings, given the extensive robustness analysis further on in this paper. The proposed AHP weights are then converted on a MACBETH scale using linear programming. This approach also circumvents potential compatibility problems, as MACBETH requires a very high level of consistency in the judgment table to be able to calculate the attractiveness levels for all the alternatives. From the performance table of the alternatives and the MACBETH weights calculated from the AHP approach, the overall attractiveness of the alternatives can be calculated. This measure of attractiveness can be interpreted as the SRI design process quality score.

In a final step sensitivity and robustness checks can again be performed. The focus is primarily on the weight of criteria and the impact on the final scores. Both the weights from a second expert and stability intervals are used to perform this robustness check. We implement the MACBETH methodology using the M-MACBETH software package.

2.4.3. Multi-attribute utility theory

Multi-attribute utility theory (MAUT) is a value measurement method that distinguishes itself by the use of utility functions to transform the performance table (Keeney & Raiffa, 1976). The analysis starts again from a hierarchy of criteria and the scoring of considered alternatives with respect to these criteria. Next, the marginal scores for every criterion are transformed into utilities using marginal utility functions that need to be elicited. This is typically the most important and difficult step in MAUT. To find the marginal utility function for every criterion, a lot of information is needed from the decision maker. Given our definition of criteria as true or false statements, however, this elicitation process is simplified to a great extent. Since every criterion can only receive one of two scores (1 or 0), the utility function is a simple two-point discrete function that needs no further information to elicit.

Next to the elicitation of utility functions, the decision maker needs to specify the weights for the criteria. There is no built-in way of doing this in MAUT, and thus alternatives need to be

considered. We again use the weights found from the AHP questionnaire and further examine robustness in section 4. Once the marginal utility functions and weights are established, aggregation is needed to find the global utility scores by which the alternatives can be ranked. Typically, an additive aggregation model is used, although alternatives are also possible (e.g. geometric). The obtained scores serve as the proxy for the SRI process quality of the mutual funds. Note that the weighted sum approach is nested within MAUT when all marginal utility functions are assumed to be linear.

Because of the specific definition of our criteria as individual statements that are evaluated dichotomously, the added value of MAUT in using underlying utility functions disappears. The dummy-coded criteria require a simple discrete utility function, which yields a utility table that looks exactly like the original performance table. Since the weighting process is also not specific to MAUT, but taken from the AHP approach, results found by means of MAUT are exactly the same as for AHP. Consequently, we do not include the results from MAUT in the next sections, as in our specific case, MAUT is a special case of AHP.

2.4.4. PROMETHEE

PROMETHEE stands for “Preference Ranking Organization METHod for Enriched Evaluation” and belongs to the outranking approach. It was originally developed by Brans and Vincke (1985) and can be somewhat considered as an improvement over the ELECTRE outranking techniques that originated with Roy (1985). A hierarchy of criteria, the alternatives considered to the problem and their performances are again required at the start of the process. Then, preference degrees are computed by means of a preference function. A preference degree expresses how the decision maker prefers one alternative to another for a certain criterion, based on an underlying preference function that needs to be decided on. This kind of unicriterion performance comparison between alternatives is typical for the outranking approach. The elicitation of the preference function is again straightforward. Because of the definition of the criteria, the preference function will be discrete and the preference degrees will be exactly the same as the scores from the performance table. However, further calculations in PROMETHEE add value and discrimination to the indicator scores with respect to AHP/MAUT. Given the preference

degrees, unicriterion flows are computed, which measure how every alternative is preferred to other alternatives. Specifically, PROMETHEE II is used to construct a complete ranking of the alternatives. The obtained net flows can be used as a quantitative input to control for differences in social responsibility, just like for the value measurement methods considered in our research.

As for MAUT, there is no built-in way of defining the weights of the different criteria. Therefore, we again implement weights obtained from the AHP questionnaire, which can then be used to aggregate the unicriterion flows to global net flows per alternative. These global net flows are the quantitative output from PROMETHEE that can be used as a proxy for the process quality indicator.

An interesting feature of PROMETHEE is that it can be used in conjunction with the GAIA (Geometric Analysis for Interactive Aid) visualization procedure, which presents the results in a two- or three-dimensional plane. GAIA not only presents results, but can also be used to enrich the sensitivity analysis to test for the quality of the obtained results. The PROMETHEE method is implemented by means of the D-Sight software.

2.4.5. TOPSIS

TOPSIS is the acronym for “Technique of Order Preference Similarity to the Ideal Solution” and is part of the third class of MCDA methods, i.e. goal, aspiration and reference-level approaches (Behzadian, Otaghsara, Yazdani, & Ignatius, 2012). As for every other method, the technique starts with defining alternatives and criteria, and introducing the performance table. The only subjective parameter needed in TOPSIS is the elicitation of the weights, for which there is no built-in method. Again, we implement weights found from the experts through the AHP questionnaire. From a weighted version of the performance table, TOPSIS constructs a ranking based on the ratio of the Euclidean distances of every alternative to the ideal (true statement for every criterion) and anti-ideal solution (false statement for every criterion). This ratio is a normalized score between 0 and 1, which can be interpreted as a process quality score. The best alternative will be the one that maximizes the distance from the anti-ideal solution and minimizes the distance from the ideal solution at the same time. Changing the weights on the different

criteria can help determine the robustness of the ranking and results. We implement the TOPSIS analysis by simply programming the method in a spreadsheet.

Most of the methods based on a goal programming, aspiration or reference level approach are not suited for the purpose of our process quality indicator. By including TOPSIS, we include methods from all three schools of MCDA, which greatly contributes to the representativeness of our search for the best performing SRI process indicator.

3. Results

For each of the four considered methods the same routine is used to obtain scores on the process-oriented social responsibility of a mutual fund. We first introduce the criteria and the alternatives, which were presented earlier. Next, the performance table (Appendix B), combining criteria to the different alternatives, is filled in by examining publicly disclosed fund information (e.g. prospectus and transparency documents). Then we introduce the weights found from one of the experts filling in the AHP-type questionnaire. In the next section we consider the weights found from the other expert as a way of examining robustness of results. The performance table, together with the weights, is then transformed by means of the four specific MCDA methods. From this transformation, we find our process-oriented sustainability scores, which are presented in Table 4.

Summary results (non-normalized)

<i>Alternatives</i>	<i>AHP</i>	<i>PROMETHEE</i>	<i>TOPSIS</i>	<i>MACBETH</i>
Max	0.158	0.7016	0.082050188	1
Triodos Sustainable Pioneer Fund	0.15	0.6739	0.08144545	0.9336
KBC EcoFund World	0.139	0.6376	0.07966546	0.8622
Dexia L Sustainable World	0.138	0.632	0.07963925	0.852
ING L Invest Sustainable Equity	0.128	0.6167	0.079327669	0.8061
BNPP L1 Equity World Aqua	0.125	0.5861	0.078209529	0.7143
BNPP L1 Equity World	0.052	0.3295	0.030019509	0.4336
ING L Invest World	0.049	0.3175	0.029618428	0.3979
KBC Equity Fund World	0.032	0.2585	0.022141209	0.2857
Dexia Quant Equities World	0.029	0.2466	0.021608729	0.25

Summary results (normalized)

<i>Alternatives</i>	<i>AHP</i>	<i>PROMETHEE</i>	<i>TOPSIS</i>	<i>MACBETH</i>
Max	100.00%	100.00%	100.00%	100.00%
Triodos Sustainable Pioneer Fund	94.94%	96.05%	99.26%	93.36%
KBC EcoFund World	87.97%	90.88%	97.09%	86.22%
Dexia L Sustainable World	87.34%	90.08%	97.06%	85.20%
ING L Invest Sustainable Equity	81.01%	87.90%	96.68%	80.61%
BNPP L1 Equity World Aqua	79.11%	83.54%	95.32%	71.43%
BNPP L1 Equity World	32.91%	46.96%	36.59%	43.36%
ING L Invest World	31.01%	45.25%	36.10%	39.79%
KBC Equity Fund World	20.25%	36.84%	26.98%	28.57%
Dexia Quant Equities World	18.35%	35.15%	26.34%	25.00%

Table 4: Summary of results

The top panel presents the results as they are found from the different methodologies. We included one theoretical alternative that indicates the maximum attainable score. The bottom panel presents normalized results with respect to the theoretical maximum.

In the top panel of Table 4 we present the scores as we find them from the different methods. A first important observation is that the ranking of the funds is completely consistent across the four methodologies. On top of the ranking we find the theoretical alternative for which the performance table is maximized, i.e. the alternative that receives a “true” statement for every criterion. This theoretical alternative is what other funds should consider in order to improve their current social responsibility score. The fund by Triodos is found to be the most sustainable alternative from a process point of view. This result should not be surprising, as Triodos is a niche player in the banking industry focusing specifically on sustainability and social responsibility. The rest of the ranking is also in line with generally accepted intuition in the Belgian financial industry. Behind Triodos, we first find the SRI funds by the four traditional banks active on the Belgian market, followed by the conventional funds. An interesting

observation is that the ranking among the traditional banks is inverted going from SRI to conventional funds. For SRI funds, KBC and Dexia seem to have the most thorough design methodology. However, BNP Paribas Fortis and ING are able to better design their conventional funds with respect to social responsibility, in contrast to KBC and Dexia.

One issue with the scoring output from the four methodologies is the difference in scaling. Therefore, we present the normalized scores in the bottom panel of Table 4. Scores are obtained by dividing the original score by the score of the theoretical optimal alternative for the respective methodology. With respect to ranking, the exact same insights are found as for the top panel of Table 4. Additionally, we now get some grasp of the relative differences in scores between the alternatives over the different methodologies. We notice that the Triodos fund scores relatively well and is very close to the theoretical maximum. However, the difference between the ideal and the Triodos fund somewhat differs between the methods. From the TOPSIS-ranking, the Triodos fund is less than one percent away from the optimum. For the other methods, this difference is slightly larger. The same kind of relative differences are found for the remainder of the ranking as well. Apart from some variation in relative performance, general insights and ranking of the alternatives is completely consistent over the four methodologies.

As indicated before, the scores found from the different methods can be instrumental in a number of ways. First, a ranking can be made like in Table 4, which tells us something about the relative performance with respect to social responsibility of different funds provided by financial institutions. Second, the scores can be used in academic applications as an extra quantitative input in performance research (e.g. asset pricing regressions, DEA). Finally, financial institutions can implement the scores and the applied methodologies in a professional context to improve the design process of their funds with respect to social responsibility.

4. Discussion

From the available set of MCDA methods and the three schools of thought, we implemented four eligible methodologies. Now, we aim to determine which specific method is most appropriate as a process quality indicator. For every application, we require that the obtained scores are as

robust as possible. Other criteria depend on the type of application. First, we consider the application of the obtained scores to further examine the performance of SRI funds (academic application). To this end, we also take into account ease of implementation and understanding of the methodology. Second, we also examine what methodology is most recommended to help banks optimize the design of a fund with respect to social responsibility (professional application). Here, we consider the extent to which sensitivity analyses can be performed. A summary of these considerations can be found in the application comparison table (Table 5).

Application comparisons	AHP	MACBETH	PROMETHEE	TOPSIS
Overall robustness		x	x	x
	<i>Academic application</i>			
Ease of implementation	x		x	x
Ease of understanding	x		x	x
	<i>Professional application</i>			
Extensiveness of sensitivity analysis			x	

Table 5: Application comparison table

Comparison of the applied methods with respect to the criteria relevant to both the academic and professional application of the indicator.

Overall robustness of results is obviously an important criterion for both types of application. To comment on robustness, we calculate stability intervals, which indicate between which extreme values the weights of a particular criterion can change such that the best alternative remains on top of the ranking. Also, we consider the weights found from a second independent SRI expert. Stability intervals and weights for each criterion with respect to the four considered methods are presented in Table 6.

For every criterion, we first indicate the allocated weight given the responses of the first and second expert on the AHP questionnaire. Next, we present the lower- and upper bound between which criteria weights can change in order for the best alternative to remain on top of the ranking. Note that for MACBETH the AHP weights are converted on the MACBETH interval scale using linear programming. Therefore, we included the transformed MACBETH weights as well in two separate columns next to the MACBETH stability intervals. For example, the first alternative receives a weight of 33.15% from the first expert and 2.40% from the second expert (8.67% vs. 3.48% for the MACBETH weights). Apparently, there is a strong lack of consensus

between the experts with regard to this criterion. However, the stability interval for the AHP method goes from 0% to 95.5%; the other methods have a complete stability interval between 0% and 100%. Concretely, this means that the fund from Triodos bank will remain the best performing fund when the weight allocated to the first criterion changes between 0% and 95.5% (resp. 100%). If the weight should become 95.6% or higher, the KBC SRI fund will come on top of the ranking for the AHP methodology. For the other methodologies, the top of the ranking will be insensitive to changes in the weight of the first criterion. To summarize the information from the numerous stability intervals for every methodology across the 20 criteria, we calculate the average distance between the lower- and upper bound. AHP seems to be least robust with an average interval width of 89.87%, which is clearly lower than for the other methodologies. PROMETHEE and TOPSIS have an average interval width of 95.35%. For MACBETH, the average width is 95.38%. Note that we discard this very small differential between PROMETHEE and TOPSIS, and MACBETH. Overall, we could conclude that AHP clearly underperforms the other methodologies with respect to robustness. PROMETHEE, TOPSIS and MACBETH can be considered “most robust” and thus meet the first criterion for both the academic and professional application of the indicator.

If we look at the stability intervals for PROMETHEE, TOPSIS and MACBETH in more detail, we see that the interval width is 100% for every criterion, except one. Apparently, the weight on the criterion determining whether the fund invests principally in companies that invest in sustainable technologies is really driving results. However, considering the judgments from our two individually consulted experts, there seems to be no impact on the ranking whatsoever. It might be of interest to further investigate the role and importance of this particular criterion with a panel of SRI experts representing different stakeholders from the field. This, however, falls beyond the scope of the current paper, but presents an interesting alley for future research.

In addition to the average interval width, we also calculated the average absolute distance between the weights from the first and the second expert in the first (and fifth) summary column. We find this distance to be 4.97% (2.81% for the MACBETH weights), which is quite significant considering that an equal-weighting of the criteria would lead to individual weights of 5%. The average distance is thus almost 100% of the equal weight. Looking at the differences in more

detail, we note several strong deviations, particularly for the first and second criterion. These observations confirm the complexity of the debate on relevant criteria and associated weights to capture social responsibility. It is therefore doubtful whether large SRI expert panels would lead to a useful compromise. Nevertheless, despite this apparent lack in consensus between the two experts, the top of the ranking does not change between the two sets of weights. This confirms earlier research finding that the weights of criteria do not matter all that much. Consequently, we can validate our approach, involving two independent SRI experts and an extensive sensitivity analysis.

Bottom-level criteria weight robustness

Criterion	Weight E1	Weight E2	AHP		PROMETHEE		TOPSIS		Weight E1	Weight E2	MACBETH	
			L	U	L	U	L	U			L	U
Priority screening process	33.15%	2.40%	0.00%	95.50%	0.00%	100.00%	0.00%	100.00%	8.67%	3.48%	0.00%	100.00%
Data gathering and analysis of sustainability by independent external specialists (e.g. EIRIS)	2.16%	23.61%	0.00%	95.50%	0.00%	100.00%	0.00%	100.00%	3.57%	9.45%	0.00%	100.00%
Incorporation of SRI principles established by reputable organizations (e.g. UN SRI, Febelfin)	5.03%	3.27%	0.00%	93.80%	0.00%	100.00%	0.00%	100.00%	6.12%	4.48%	0.00%	100.00%
Information from stakeholders and relevant NGOs	0.56%	6.80%	0.00%	94.90%	0.00%	100.00%	0.00%	100.00%	2.04%	7.96%	0.00%	100.00%
Best-in-class approach for criteria with respect to ESG criteria	5.74%	9.20%	0.00%	94.10%	0.00%	100.00%	0.00%	100.00%	7.65%	8.95%	0.00%	100.00%
Use of sector specific positive criteria	2.76%	2.13%	0.00%	94.40%	0.00%	100.00%	0.00%	100.00%	4.59%	2.99%	0.00%	100.00%
Investment is principally (> 75%) in companies that invest in sustainable technologies	0.49%	4.43%	0.00%	3.00%	0.00%	7.08%	0.00%	7.08%	1.02%	5.97%	0.00%	7.62%
Use of categorical rejects	7.57%	2.47%	0.00%	93.90%	0.00%	100.00%	0.00%	100.00%	8.16%	3.98%	0.00%	100.00%
Assessment by means of negative criteria	7.57%	5.14%	0.00%	95.60%	0.00%	100.00%	0.00%	100.00%	8.16%	6.47%	0.00%	100.00%
Exclusion of unsustainable technologies	7.57%	3.56%	6.40%	100.00%	0.00%	100.00%	0.00%	100.00%	8.16%	4.98%	0.00%	100.00%
A research teams checks legal and regulatory developments, trends and behavior of companies such that criteria are in line with recent societal developments	5.25%	1.97%	0.00%	95.60%	0.00%	100.00%	0.00%	100.00%	6.63%	2.49%	0.00%	100.00%
Monitoring if portfolio is consistent with defined criteria (continuously, sector specific or occasion specific)	1.75%	5.91%	0.00%	93.20%	0.00%	100.00%	0.00%	100.00%	3.06%	6.96%	0.00%	100.00%
Companies are informed about conclusions selection methodology	0.24%	1.93%	0.00%	93.60%	0.00%	100.00%	0.00%	100.00%	0.51%	1.99%	0.00%	100.00%
Active engagement policy (constructive and critical dialogue with companies in portfolio)	0.51%	0.93%	0.00%	93.50%	0.00%	100.00%	0.00%	100.00%	1.53%	0.50%	0.00%	100.00%
Active voting policy (voting at companies' shareholder meetings)	2.83%	1.34%	0.00%	93.60%	0.00%	100.00%	0.00%	100.00%	5.11%	1.00%	0.00%	100.00%
Release of qualitative information about the screening process (e.g. applied screens)	4.00%	3.90%	0.00%	93.30%	0.00%	100.00%	0.00%	100.00%	5.62%	5.47%	0.00%	100.00%
Release of quantitative information about the screening process (e.g. scores)	4.00%	3.90%	0.00%	100.00%	0.00%	100.00%	0.00%	100.00%	5.62%	5.47%	0.00%	100.00%
Release of current portfolio	2.23%	9.13%	0.00%	93.60%	0.00%	100.00%	0.00%	100.00%	4.08%	8.45%	0.00%	100.00%
Compliance with external transparency guidelines (e.g. Eurosif/Belsif)	0.91%	1.80%	0.00%	93.20%	0.00%	100.00%	0.00%	100.00%	2.56%	1.50%	0.00%	100.00%
Board of experts	5.56%	6.25%	0.00%	93.50%	0.00%	100.00%	0.00%	100.00%	7.14%	7.46%	0.00%	100.00%
Summary	4.97%		89.87%		95.35%		95.35%		2.81%		95.38%	

Table 6: Stability intervals

Presentation of the stability intervals, between lower-bound (L) and upper-bound (U), for each of the 20 bottom-level criteria with respect to the four applied methods. The average length of the intervals for every method is calculated on the bottom of the table, as well as the average absolute distance between weights from the two experts (bottom of column 1 and 4).

From the robustness analysis, we learn that AHP underperforms other eligible MCDA methods. This observation is important, given that earlier research on social responsibility indicators has only implemented the AHP methodology (Pérez-Gladish & M'Zali, 2010; De Moor et al., 2012). Given that three methods prove to be equally robust, we need to consider some additional criteria to determine what method is best for calculating process-oriented social responsibility scores in future research. For the academic application, we first take into account ease of implementation. In order for the method to be feasible for future research in financial economics, it needs to be easy to implement using either existing software packages or by simple programming in spreadsheet or statistical software. Ideally, the software should be available for free to academia. MACBETH is only available upon purchasing the associated MACBETH-M software, which complicates implementation in future research. Some good PROMETHEE software packages are available for free (e.g. Visual PROMETHEE), although we used a costly alternative called D-Sight. With TOPSIS, there is no need for software, as the method can easily be programmed into a spreadsheet program like Microsoft Excel or a statistical software package like R. We also take into account the ease of understanding and transparency of the different methods. When implementing the obtained scores in mutual fund performance research, it is desirable that it is very clear how the scores were calculated. A conceptually transparent and intuitive methodology is therefore preferred. We consider both TOPSIS and PROMETHEE to be sufficiently transparent and easy to understand. The MACBETH methodology is more complicated and black box and thus less suited for implementation in further financial research. Taking the proposed criteria together, both PROMETHEE and TOPSIS are eligible for implementation in future research on the relationship between social responsibility and financial performance. Note, however, that apart from this particular application of MCDA methodologies, PROMETHEE is considered more advanced and less prone to methodological issues under a variety of data characteristics. TOPSIS is known to sometimes lead to illogical results, especially when an alternative has extreme performances on different criteria. Even though such TOPSIS methodological issue was not found in our specific application, we recommend the use of PROMETHEE in future academic applications. PROMETHEE proves to be highly robust, easy to implement and understand, and can also be considered least prone to issues under a variety of possible performance characteristics.

Another application of the scores can be found in a more professional context within the asset management and fund industry. For fund managers, it would be interesting to have a tool that provides the possibility to gain more insights into how scores were calculated and how the design of a new fund could be optimized with respect to social responsibility. Practically, such insights can be obtained by means of a thorough sensitivity analysis. To determine the most recommended methodology for professional applications, we therefore take into account the extensiveness of the available sensitivity analysis tools. Again, we find PROMETHEE to be the most appropriate methodology from the set of robust options. As explained earlier, PROMETHEE can be used in combination with a visual interactive method called GAIA, which greatly adds insights into how a score was obtained and what the best actions would be to improve it. Of most interest is the so-called GAIA plane, which provides insights into the importance of the different criteria, irrespective of the allocated weights. For example, it could very well be that a highly weighted criterion ends up being relatively unimportant in the calculation of the score and the construction of the ranking. Alternatively, low-weighted criteria might be relatively important in the eventual analysis. These kinds of insights can help a fund manager to efficiently detect opportunities for optimizing the design of a fund with respect to social responsibility. Both D-Sight and Visual PROMETHEE provide the software to perform such analyses.

The M-MACBETH software also comes with a built-in sensitivity procedure, although it is less straightforward to determine stability intervals and to gain insights into how the design of a fund can be improved. The MACBETH sensitivity tools are also less graphical, which makes the sensitivity analysis less intuitive for fund managers to use. As TOPSIS is simply programmed into a spreadsheet or statistical software, there is no standard or graphical way of performing sensitivity analyses.

5. Conclusion

In our paper we develop a process-oriented social responsibility indicator for mutual funds using multi-criteria decision analysis. This indicator is intended for further application in both academic and professional settings, rather than for individual investor purposes. Upon comparing

the range of possible MCDA methods, we find AHP, PROMETHEE, MACBETH and TOPSIS to be eligible for our problem. Twenty criteria are established from the literature, and are combined with an illustrating sample of nine mutual funds available on the Belgian market. Consulting two experts based on an AHP-type questionnaire, we establish relative criteria priorities. We design the performance table in such a way that it can be completed using only publicly disclosed fund information, without any further need for expert judgment. Using the four different MCDA methods, we transform the performance and the weight table into scores and a ranking, which provides more insight into the social responsibility of a mutual fund from the process and design perspective.

The obtained ranking is perfectly consistent across the four considered methods and is also in line with generally accepted intuition in the industry. From the stability intervals and the weights from a second expert, we find the results for PROMETHEE, MACBETH and TOPSIS to be highly robust. We consider two possible applications of the social responsibility scores. First, the scores can be used in academic research to gain more nuanced insights into the performance of SRI funds. Both parametric (e.g. asset pricing; Bauer, Koedijk, & Otten, 2005) and non-parametric (e.g. DEA; Basso & Funari, 2003) methods can draw from the scores to control for differences in process-oriented social responsibility to gain a better understanding of the relationship between return, risk and sustainability. Given its robustness, the ease to both understand and implement the method and its performance under a wide variety of data characteristics, we find PROMETHEE to be the most recommended methodology for academic applications. Second, fund managers can use the MCDA indicator and the scores in a professional setting to optimize the design of newly launched mutual funds with respect to social responsibility. Again, PROMETHEE seems most appropriate, given its extensive range of built-in possibilities to assess sensitivity of obtained results, including the GAIA visual interactive method.

Our paper presents an innovative approach to the design of a process-oriented social responsibility indicator for mutual funds, and determines that PROMETHEE is the recommended methodology for both future academic and professional applications. We contribute to the existing literature by considering the entire set of established MCDA methods,

instead of just AHP (Pérez-Gladish & M’Zali, 2010; De Moor et al., 2012) and by determining what the best methodology is for both future academic and professional applications. Additionally, we design the indicator in such a way that it is feasible to implement in future academic and professional applications, without need for elaborate expert or individual judgment, which is in contrast to the different attempts to develop a SRI framework for individual investors (e.g. Hallerbach et al., 2004; Cabello et al., 2014). Future research could have a closer look at the consistent family of criteria and possible cross-cultural differences. Starting from our research, the PROMETHEE-based indicator can now be implemented in SRI performance research and in professional applications to optimize the design of new mutual funds with respect to social responsibility. In addition to our illustration for the Belgian market, the proposed indicator could also be applied to other markets. Finally, it could be insightful to develop a similar MCDA indicator that yields in classes of funds with respect to social responsibility, rather than quantitative scores (e.g. using the PROMETHEE-based FLOWSORT methodology). Such a tool could, for example, be helpful for regulators to help classify and label mutual funds.

Appendix A: Description of criteria

- Priority screening process: the fund first executes the screening process, after which a financial analysis is implemented (not the other way around).
- Independent data gathering and analysis of sustainability:
 - An independent external specialist company (e.g. EIRIS) gathers the necessary data and analyses sustainability
 - SRI principles established by national (e.g. Febelfin) and international (e.g. United Nations) organizations are referred to and reflected in the portfolio selection criteria
 - NGOs and relevant stakeholders are involved in the data gathering process
- Positive selection criteria:
 - A best-in-class approach (top 30% performing companies in an industry) is developed

- with respect to ESG criteria
- Sector specific criteria are used
 - Investment is principally (> 75%) in companies that actively invest in sustainable technologies (e.g. green electricity, CO₂ reducing machinery, waste reduction, water quality)
 - Negative selection criteria:
 - Categorical rejects using predefined exclusion criteria (e.g. companies involved with nuclear power, tobacco and/or weapons).
 - Contestable activities (e.g. gambling, genetically modified organisms, bio hazards) can lead to exclusion, depending on the extent of involvement and the context (more nuanced than categorical rejects).
 - No investments in unsustainable technologies, irrespective of possible social or ecological measures.
 - Unsustainable technologies: coal plants, nuclear energy, crude oil, coal to liquid, macro-scale hydro power.
 - Monitoring and updating:
 - A research team checks legal and regulatory developments, trends and behavior of companies such that criteria remain in line with recent societal developments
 - Portfolio is monitored for compliance with the set of defined criteria (continuously, sector or event specific)
 - Dialogue
 - Companies are informed about conclusions of the fund's research and get suggestions for improvement of social performance.

- There is an active engagement policy, which means that there is a constructive and critical dialogue with the companies included in the fund's portfolio in light of positively influencing corporate behavior.
- There is an active voting policy, which means that representatives of the fund attend shareholder meetings, speak up and vote to change companies' behavior for the better.
- Transparency
 - Release of qualitative information about the screening process (e.g. criteria used, description of process).
 - Release of quantitative information about the screening process (e.g. scores for individual funds, investment universe).
 - The composition of the portfolio is continuously disclosed.
 - The fund complies with national and international transparency guidelines (e.g. Eurosif/Belsif).
- Board of experts: a board of experts is consulted to help develop the methodology for building the portfolio.

Appendix B: Performance table

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
BNPP S	1	1	1	0	0	0	1	1	1	0	1	1	0	1	1	1	0	1	1	1
Dexia S	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1
ING S	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	1
KBC S	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1
Triodos	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1
BNPP C	0	1	1	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	0	0
Dexia C	0	0	1	0	0	0	0	1	0	0	0	0	0	1	1	0	0	1	0	0
ING C	0	0	1	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	0	0
KBC C	0	1	1	0	0	0	0	1	0	0	0	0	0	1	1	0	0	1	0	0

BNPP S	BNPP L1 Equity World Aqua
Dexia S	Dexia L Sustainable World
ING S	ING L Invest Sustainable Equity
KBC S	KBC EcoFund World
Triodos	Triodos Sustainable Pioneer Fund
BNPP C	BNPP L1 Equity World
Dexia C	Dexia Quant Equities World
ING C	ING L Invest World
KBC C	KBC Equity Fund World

C1	Priority screening process
C2	Data gathering and analysis of sustainability by independent external specialists (e.g. EIRIS)
C3	Incorporation of SRI principles established by reputable organizations (e.g. UN SRI, Febelfin)
C4	Information from stakeholders and relevant NGOs
C5	Best-in-class approach for criteria with respect to ESG criteria
C6	Use of sector specific positive criteria
C7	Investment is principally (> 75%) in companies that invest in sustainable technologies
C8	Use of categorical rejects
C9	Assessment by means of negative criteria
C10	Exclusion of unsustainable technologies
C11	A research team checks legal and regulatory developments, trends and behavior of companies such that criteria are in line with recent societal developments
C12	Monitoring if portfolio is consistent with defined criteria (continuously, sector specific or occasion specific)
C13	Companies are informed about conclusions selection methodology
C14	Active engagement policy (constructive and critical dialogue with companies in portfolio)
C15	Active voting policy (voting at companies' shareholder meetings)
C16	Release of qualitative information about the screening process (e.g. applied screens)
C17	Release of quantitative information about the screening process (e.g. scores)
C18	Release of current portfolio
C19	Compliance with external transparency guidelines (e.g. Eurosif/Belsif)
C20	Board of experts

References

- Bana e Costa, C., De Corte, J.-M., & Vansnick, J.-C. (2005). On the mathematical foundation of MACBETH. In J. Figueira, S. Greco, & M. Ehrgott (Eds.), *Multiple criteria decision analysis: State of the art surveys* (pp. 409–437). New York: Springer-Verlag.
- Basso, A., & Funari, S. (2003). Measuring the performance of ethical mutual funds: A DEA approach. *Journal of the Operational Research Society*, 54(5), 521–531.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751–1767.
- Behzadian, M., Otaghsara, K., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*, 39(17), 13051–13069.
- Belton, V., & Stewart, T. J. (2002). *Multiple criteria decision analysis: An integrated approach*. Boston: Kluwer Academic Publishers.
- Brans, J. P., & Vincke, P. (1985). A preference ranking organisation method: the PROMETHEE method for multiple criteria decision-making. *Management Science*, 31(6), 647–656.
- Cabello, J. M., Ruiz, F., Pérez-Gladish, B., & Méndez-Rodríguez, P. (2014). Synthetic indicators of mutual funds' environmental responsibility: An application of the Reference Point Method. *European Journal of Operational Research*, forthcoming.
- De Moor, L., Devooght, K., & De Bondt, C. (2012). Een rangschikking volgens het duurzaam en maatschappelijk verantwoord karakter van vier Belgische beleggingsfondsen (A ranking of four Belgian mutual funds according to their sustainable and socially responsible character). *Bank- en Financiewezen*, 76(2-3), 161–167.
- European Sustainable Investment Forum (Eurosif). (2010). *European SRI study*. Available at: <http://www.eurosif.org>.
- European Sustainable Investment Forum (Eurosif). (2012). *European SRI Study*. Available at: <http://www.eurosif.org>.

- Febelfin. (2012). *Aanbeveling inzake duurzame financiële producten (Recommendation concerning sustainable financial products)*. Available at <http://www.febelfin.be>.
- Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multiple criteria decision analysis: State of the art surveys*. New York: Springer-Verlag.
- Forum for Sustainable and Responsible Investment. (2012). *Sustainable and responsible investing trends in the US*. Available at <http://www.ussif.org>.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman/Ballinger.
- Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. *New York Times Magazine*, pp. 32–33, 122, 124, 126.
- Hallerbach, W., Ning, H., Soppe, A., & Spronk, J. (2004). A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research*, 153(2), 517–529.
- Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: John Wiley & Sons.
- KPMG, & ALFI. (2013). *European Responsible Investing Fund Survey 2013*. Available at <http://www.kpmg.com/LU>.
- Mackey, A., Mackey, T. B., & Barney, J. B. (2007). Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Academy of Management Review*, 32(3), 817–835.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- Moskowitz, M. (1972). Choosing socially responsible stocks. *Business and Society Review*, 1, 71–75.
- Pérez-Gladish, B., & M'Zali, B. (2010). An AHP-based approach to mutual funds' social performance measurement. *International Journal of Multicriteria Decision Making*, 1(1), 103.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008a). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302–322.

- Renneboog, L., Ter Horst, J., & Zhang, C. (2008b). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, 32(9), 1723–1742.
- Roy, B. (1974). Critères multiples et modélisation des préférences: L'apport des relations de surclassement (Multiple criteria and modeling of preferences: Input of outranking relations). *Revue d'Economie Politique*, 1, 1–44.
- Roy, B. (1985). *Méthodologie multicritère d'aide à la décision (Multicriteria decision aid)*. Paris: Economica.
- Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3), 234–281.
- Saaty, T. L. (1980). Analytic hierarchy process. *Decision Analysis*, 50(1), 579–606.
- Sparkes, R., & Cowton, C. J. (2004). The maturing of socially responsible investment: A review of the developing link with corporate social responsibility. *Journal of Business Ethics*, 52(1), 45–57.
- United Nations. (2013). The six principles. *Principles for Responsible Investment*. Available at <http://www.unpri.org>